

Seminario di Cultura Digitale

Università di Pisa

9 marzo 2016



Alla ricerca dell'antenato comune

Sommario

- ▶ Antenato comune più recente (*Most Recent Common Ancestor – MRCA*)
- ▶ Utilizzo di simulazioni informatiche (*in silico*)
- ▶ Digitalizzazione delle genealogie
 - ▶ Banche dati online di genealogie nobiliari
 - ▶ Genealogie online come social network

Modello



Dati

Genealogie

- ▶ Genetica
 - ▶ Trasmissione del patrimonio genetico
- ▶ Genealogia
 - ▶ Processo (di *branching*) genitore-figlio

Genetica

Genealogia

deltaG





Ripetizione degli antenati

- ▶ E' un fenomeno intuitivo: gli antenati di ciascuno crescono in modo esponenziale e supererebbero in breve l'intera popolazione mondiale
- ▶ Derrida et al. (1999), *Statistical Properties of Genealogical Trees*: non ci occuperemo specificamente di questo ma di un fenomeno collegato: la coalescenza degli antenati
- ▶ Ricerca degli Antenati Comuni (CA) e, tra questi, il più recente (MRCA)
- ▶ Ricerca dell'Identical Ancestor Point (IAP): momento in cui tutte le persone in vita all'epoca o sono antenati comuni di tutti o non lo sono di nessuno



Antenato comune genetico

Si prende in considerazione materiale genetico non soggetto a ricombinazione

➤ **Linea materna**

➤ DNA mitocondriale

Eva mitocondriale

100000 – 200000 anni fa

Cann et al. (1987), *Mitochondrial DNA and human evolution*

➤ **Linea paterna**

➤ Cromosoma Y

Adamo Y-cromosomico

270000 anni fa

Pääbo (1995), *The Y chromosome and the origin of all of us (man)*

Antenato comune genealogico (I)

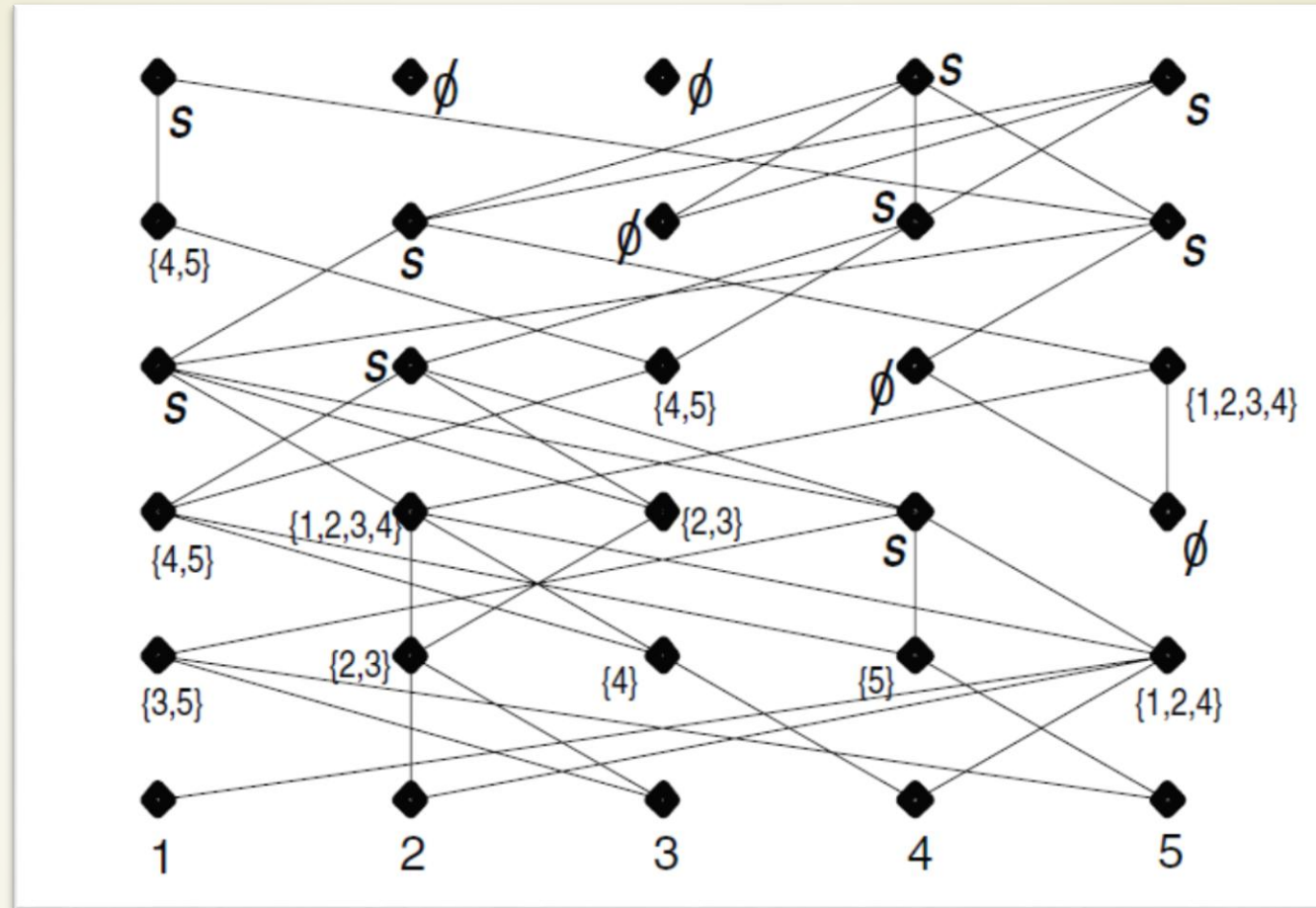
Chang (1999), *Recent common ancestors of all present-day individuals*

- ▶ Modello biparentale
- ▶ Popolazione costante
- ▶ Accoppiamento casuale (due scelte con uguale probabilità)
- ▶ Nessuna distinzione di sesso

Teorema 1: Il numero di generazioni che separa una popolazione di dimensioni n dal MRCA tende a $\log_2(n)$ quando n tende ad infinito

Teorema 2: Lo IAP si verifica a $1,77\log_2(n)$ generazioni rispetto alla popolazione di partenza

Il modello non è realistico ma i risultati non sono lontani dal vero



Esempio del modello di Chang

CA: gen. 2 – IAP: gen. 5

Antenato comune genealogico (II)

Rohde et al. (2004), *Modelling the recent common ancestry of all living humans*

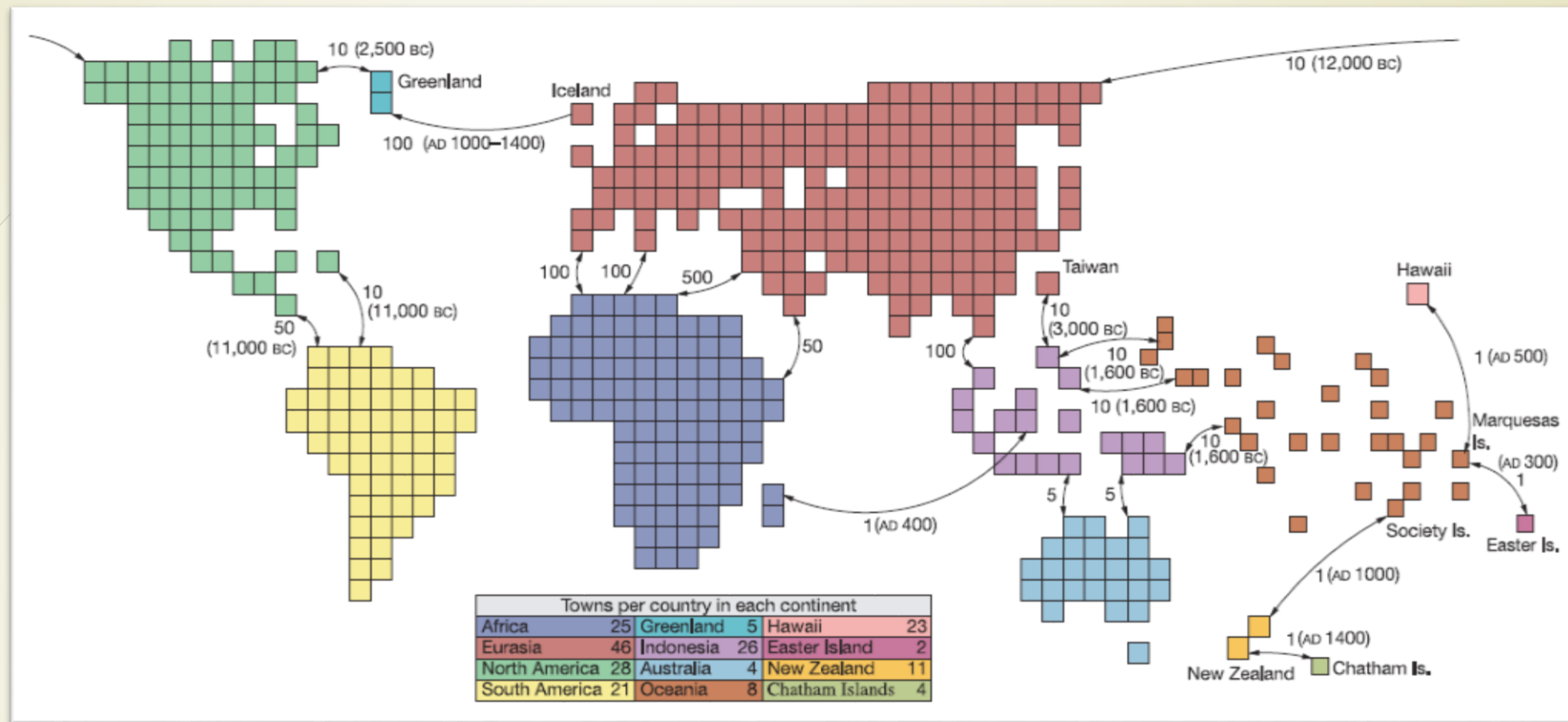
- ▶ Simula la popolazione mondiale dal 20000 a.C. al 2000 d.C.
- ▶ Due simulazioni: una con modello probabilistico, una con metodo Monte Carlo
- ▶ Accoppiamento non casuale ma legato a prossimità, geografia, cultura, lingua, classe sociale
- ▶ Prevede continenti, paesi, città, porti e individui migranti
- ▶ Le generazioni si sovrappongono simulando le «età della vita»

MRCA: 1415 a.C. – 55 d.C. **IAP:** 5353 a.C. – 2158 a.C.

Risultati compatibili con i teoremi di Chang

60GB di dati compressi per esperimento derivanti da 3 ore di elaborazione

Dalle 5 alle 10 ore per individuare il MRCA (Pentium 4 2,7GHz – 2GB RAM)



Il mondo secondo Rohde

12 continenti, 497 paesi, 15059 città, 60 milioni di sims

Dal modello ai dati



Sono in massima parte digitalizzate. Due sono i siti di riferimento:

- **Genealogics** (www.genealogics.org): 680.000 schede (gen. 2016)
- **Roglo** (roglo.eu/roglo): 6.327.160 schede (mar. 2016)

Occorre tuttavia **estrarre, integrare e disambiguare** i dati.



Il progetto Hochadelsdorf

Gli obiettivi

- Identificare MRCA e IAP sia dell'intero campione sia dei soli Hoch
- Studiare la distribuzione e le ripetizioni degli antenati
- Altre ricerche (cognomi, cladistica, isonimia e consanguineità, ...)

La base dati

- 16 generazioni genealogiche
- 48 individui nati tra il 1865 e il 1917
- 200.000 ca. individui identificati
- 3.145.680 soggetti (e le loro relazioni)

Lo scenario (stati sovrani al 1914)

Stato	Famiglia
Baviera	Wittelsbach
Prussia	Hohenzollern
Sassonia	Wettin
Wurttemberg	Wurttemberg
Baden	Zahringen
Hessen bei Rhein	Hessen
Hessen	Hessen
Mecklenburg-Schwerin	Mecklenburg
Mecklenburg-Strelitz	Mecklenburg
Oldenburg	Oldenburg
Sachsen-Weimar-Eisenach	Wettin
Anhalt	Ascania
Brunswick	Hanover

Stato	Famiglia
Sachsen-Altenburg	Wettin
Sachsen-Coburg-Gotha	Wettin
Sachsen-Meiningen	Wettin
Lippe	Lippe
Reuss-Gera	Reuss
Reuss-Greiz	Reuss
Schaumburg-Lippe	Lippe
Schwarzburg-Rudolstadt	Schwarzburg
Schwarzburg-Sondershausen	Schwarzburg
Waldeck-Pyrmont	Waldeck
Hohenlohe	
Isenburg	
Solms	

Paese	Famiglia
Albania	Wied
Austria-Ungheria	Lorraine
Belgio	Wettin
Bulgaria	Wettin
Danimarca	Oldenburg
Francia	Bourbon
Grecia	Oldenburg
Italia	Savoia
Lussemburgo	Nassau
Paesi Bassi	Nassau
Norvegia	Oldenburg
Portogallo	Braganca-Wettin
Regno Unito	Wettin
Romania	Hohenzollern
Russia	Oldenburg
Spagna	Bourbon

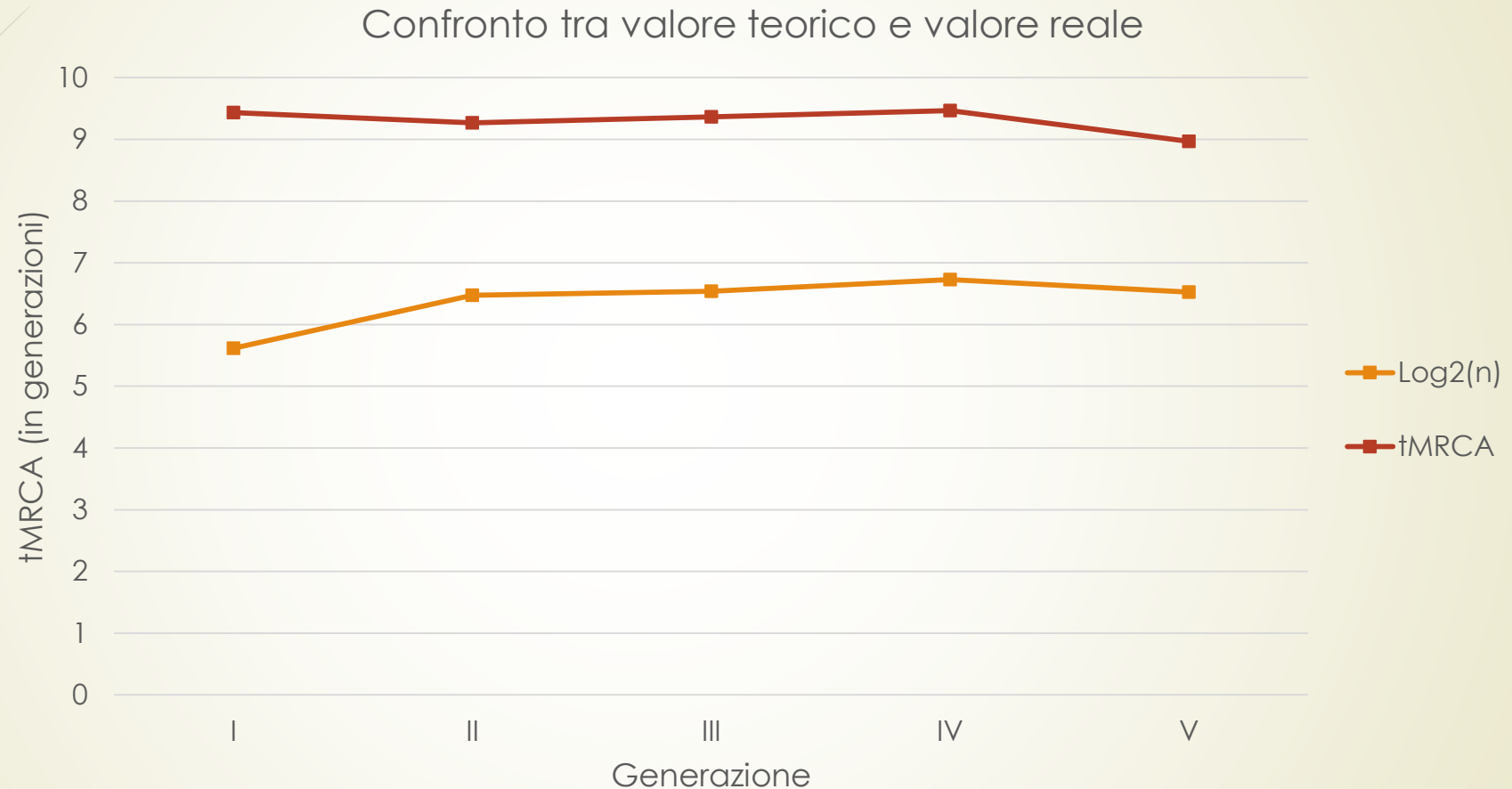
MRCA

Gen.*	Campione	Anno MRCA	tMRCA	Nome
I	49	1612	283	Joachim Ernst I Oettingen-Oettingen
II	89	1587	278	Magdalene Sybille von Preussen
III	93	1554	281	Margarethe von Schönburg-Glauchau
IV	106	1521	284	Friedrich Magnus Solms-Laubach
V	92	1506	269	Juliana Stolberg-Wernigerode

*Generazioni anagrafiche di 30 anni ciascuna secondo lo schema seguente:

I	II	III	IV	V
1880 - 1909	1850 - 1879	1820 - 1849	1790 - 1819	1760 - 1789

tMRCA



- Il MRCA si colloca a circa **9,3 generazioni** dalla popolazione di base

“ No matter the languages we speak or the colour of our skin, we share ancestors who planted rice on the banks of the Yangtze, who first domesticated horses on the steppes of the Ukraine, who hunted giant sloths in the forests of North and South America, and who laboured to build the Great Pyramid of Khufu.

”

Rohde et al. (2004), *Modelling the recent common ancestry of all living humans*

Contatti:

<http://giorgio.spugnesi.info/uni>

giorgio.spugnesi@gmail.com

Questa presentazione è coperta, per le sue parti originali, da licenza Creative Commons:
attribuzione, non commerciale, condividi allo stesso modo
<http://creativecommons.org/licenses/by-nc-sa/2.5/it/legalcode>



Bibliografia

CANN, R. L., STONEKING, M., WILSON, A. C. (1987). Mitochondrial DNA and human evolution. *Nature* 325, 31-36.

CHANG, J. T. (1999). Recent Common Ancestors of All Present-day Individuals. *Adv. Appl. Prob.* 31, 1002-1026.

DERRIDA, B., MANRUBIA, S.C., ZANETTE, D.H. (1999). Statistical Properties of Genealogical Trees. *Phys. Rev. Lett.* 82, 1987 – Published 1 March 1999.

DERRIDA, B., MANRUBIA, S.C., ZANETTE, D.H. (2000a). On the genealogy of a population of biparental individuals. *J Theor Biol.* Apr 7, 203(3), 303-15.

DERRIDA, B., MANRUBIA, S.C., ZANETTE, D.H. (2000b). Distribution of repetitions of ancestors in genealogical trees. *Physica A*, Volume 281, Issues 1-4, 15 June 2000, pages 1-16.

DORIT, R. L., AKASHI, H., GILBERT, W. (1995). Absence of polymorphism at the ZFY locus on the human Y chromosome. *Science* 268, 1183-1185.

PÄÄBO, S. (1995). The Y chromosome and the origin of all of us (man). *Science* 268, 1141-1142.

ROHDE, D. L. T., OLSON S., CHANG, J. T. (2004). Modelling the recent common ancestry of all living humans, *Nature* 431, 562-566.

Sul progetto Hochadelsdorf

Hochadelsdorf project (<http://www.df.unipi.it/~rossi/Hochadelsdorf%20project.pdf>)

ROSSI, P. (2014). Genealogie digitali (<http://www.df.unipi.it/~rossi/Genealogie%20digitali.pdf>)